

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 12:59:25 ; Search time 93 Seconds
(without alignments)
1289.456 Million cells

File: US-09-635-949-34

Perfect score: 3289

Sequence: 1 M D F I L A L V I V S S I Y I Q A A A F T G E I G L D D V S I K K G H C S E E R 582

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

100% processing. Minimum Match 98%
Maximum Match 100%

Maximum MacCn 1008
Listing first 45 summaries

Database :

- ```

1: sp:archaea:*
2: sp:bacteria:*
3: sp:bacteria:*
4: sp:fungi:*
5: sp:human:*
6: sp:invertebrate:*
7: sp:mammal:*
8: sp:muc:*
9: sp:organelle:*
10: sp:phage:*
11: sp:plant:*
12: sp:protist:*
13: sp:virus:*
14: sp:vertebrate:*
15: sp:unclassified:*
16: sp:virus:*
17: sp:bacteriap:*
18: sp:archaea:*

```

**pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query |        | DB | ID     | Description  |
|------------|--------|-------|--------|----|--------|--------------|
|            |        | Match | length |    |        |              |
| 1          | 2911   | 88.5  | 578    | 11 | Q91ZD3 | mus musculus |
| 2          | 2885.5 | 87.7  | 609    | 11 | Q923T5 | mus musculus |
| 3          | 2809.5 | 85.4  | 561    | 11 | Q91V88 | mus musculus |
| 4          | 2784   | 84.6  | 592    | 11 | Q91XL5 | mus musculus |
| 5          | 1096   | 33.3  | 554    | 4  | Q9NT67 | mus musculus |
| 6          | 1092.5 | 33.2  | 553    | 4  | Q9NZL7 | homo sapien  |
| 7          | 1085   | 33.0  | 550    | 11 | Q9JJZ5 | homo sapien  |
| 8          | 1066.5 | 32.4  | 558    | 4  | Q9UPK6 | mus musculus |
| 9          | 895    | 27.2  | 474    | 4  | Q8WY3  | homo sapien  |
| 10         | 405.5  | 12.3  | 2809   | 4  | Q96JP8 | homo sapien  |
| 11         | 402    | 12.2  | 1174   | 11 | Q93K58 | mus musculus |
| 12         | 400.5  | 12.2  | 3857   | 11 | O88840 | mus musculus |
| 13         | 396    | 12.0  | 2906   | 11 | Q9W0H9 | rattus norv  |
| 14         | 395.5  | 12.0  | 2872   | 11 | Q9W0H8 | rattus norv  |
| 15         | 385.5  | 11.7  | 608    | 11 | Q9DBE2 | mus musculus |
| 16         | 384    | 11.7  | 708    | 13 | P87363 | gallus gall  |

## ALIGNMENTS

## RESULT 1

```

Q91ZD3 PRELIMINARY; PRT: 578 AA.
Q91ZD3:
Q91ZD3:
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
Nephronectin long isoform.
NPNT OR NEPHL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NH SWISS; TISSUE=KIDNEY;
RR MEDLINE=21363579; Pubmed=11470831;
RR Brandenberger R., Schmidt A., Linton J., Wang D., Backus C., Denda S.,
RR Muller U., Reichardt L.F.;
RR *Identification and characterization of a novel extracellular matrix
RR protein nephronectin that is associated with integrin alpha8beta1 in
RR the embryonic kidney.*
RR J. Cell Biol. 154:447-458(2001).
RL -1- SIMILARITY: CONTAINS 1 HAM DOMAIN.
EMBL; AY035899; AAK96011.1; -
MCD; MGI:2148811; Npnt.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR000998; HAM_domain.
DR Pfam; PF00608; EGF; 4.
DR Pfam; PF00629; HAM; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_3.
DR PROSITE; PS00060; HAM_2; 1.
DR GlycoProtein.
KW Glycoprotein.
Q91ZD3 578 AA; 63606 MW; 4757DF54CC47DA4C CRC64;
Q91ZD3:
Q91ZD3:

```

Query Match 88.5%; Score 2911; DB 11; Length 578;  
 Best Local Similarity 88.1%; Pred. No. 5.2e-226;  
 Matches 510; Conservative 31; Mismatches 36; Indels 2; Gaps 2;

QY 1 MDFLLALVIVSSLYLQAAAEFGDGRWPRQIVSSIGLRYGGRIDCCGWARQSWGQCPFY 60  
 Db 1 MAVLLAAVLASSLYLQVAADFDGRWPRQIVSSIGLRYGGRIDCCGWARQSWGQCPFY 60

QY 61 VLKRIARIRCOLKAVCOIPKCKHGCIGPNKCKCHPGYACKTCIGVINEGGLKPRCKHR 120  
 Db 61 VLKRIARIRCOLKAVCOIPKCKHGCIGPNKCKCHPGYACKTCIGVINEGGLKPRCKHR 120

QY 121 CMNTYGSYKCYCLNGYMLMDGSCSSALTCGSMANCOYGVDDVWKGIRCCQSPSPGLQAPD 180  
 Db 121 CMNTYGSYKCYCLNGYMLMDGSCSSALTCGSMANCOYGVDDVWKGIRCCQSPSPGLQAPD 180

QY 181 GRTCDVDECATGRVSCPRFQCVNTFGSYICKCHGKFDLMYIGGYQCHIDECSLGQY 240  
 Db 181 GRTCDVDECATGRVSCPRFQCVNTFGSYICKCHGKFDLMYIGGYQCHIDECSLGQY 240

QY 241 QCSSFANCYNVRSYKCKKCKGQGGTGCYVTPKVMIPESGPVHVPKGNCTILKGDTON 300  
 Db 241 QCSSFANCYNVRSYKCKKCKGQGGTGCYVTPKVMIPESGPVHVPKGNCTILKGDTON 300

QY 301 NNWIPDVGSTWMPKPYIPPIITNRPSTKPTTRTPKPTPIPTPPPPPLPTELR-TPL 359  
 Db 301 NNWIPDVGSTWMPKPYIPPIITNRPSTKPTTRTPKPTPIPTPPPPPLPTELR-TPL 359

QY 360 PPTTPTTGLTTIAPAATSPGGITVDRVQTDQKPRGDFVIFRQPSNDLFEIFEIE 419  
 Db 360 PPTTPTTGLTTIAPAATSPGGITVDRVQTDQKPRGDFVIFRQPSNDLFEIFEIE 419

QY 420 KGVSADEAADDPCVIVHSCNFDHGLCGWIREKNDLHWEPTRDPAGGYLTVSAAKAPG 479  
 Db 420 KGVSADEAADDPCVIVHSCNFDHGLCGWIREKNDLHWEPTRDPAGGYLTVSAAKAPG 479

QY 480 GKAARLVPLGLRMHSGDLCLSPRHKVTGLHSCTLOVFKKHGAIGAAALWGRNGHGWQ 539  
 Db 480 GKAARLVPLGLRMHSGDLCLSPRHKVTGLHSCTLOVFKKHGAIGAAALWGRNGHGWQ 539

QY 540 TQTLRGADVKSIVFKGKRRGHTGEIGLDDVSLKRGRC 578  
 Db 540 TQTLRGADVKSIVFKGKRRGHTGEIGLDDVSLKRGRC 578

## RESULT 2

Q923T5 PRELIMINARY; PRT; 609 AA.

AC Q923T5;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DE Nephronectin.  
 GN NPNT OR NEPH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Brandenberger R., Schmidt A., Linton J., Hackus C., Wang D., Denda S.,  
 KA Muller U., Reichardt L.F.;  
 RT Identification and Characterization of Nephronectin, a Novel ECM  
 RT Protein that is Associated with Integrin alpha8beta1 in the Embryonic  
 RT Kidney.\*;  
 RL J. Cell Biol. 0:0-0(2001).  
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 DR EMBL: AF197008; AAK84392.1;  
 DR MGI: 2148811; Npnt.  
 DR InterPro: IPR000152; Asx-hydroxyl.  
 DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR001861; EGF\_Ca.  
 DR InterPro: IPR000998; MAM\_domain.  
 DR Pfam: PF00008; EGF\_4.  
 DR Pfam: PF00629; MAM; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_3.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_3.  
 DR PROSITE: PS01187; EGF\_CA; UNKNOWN\_3.  
 DR PROSITE: PS50060; MAM\_2; 1.  
 KW Glycoprotein.  
 SQ SQQFNCR 609 AA; 67061 MW; 44C45DE4HEEC4HMC CR764.

## Query Match 87.7%; Score 2885.5; DB 11; Length 609;

Best Local Similarity 83.6%; Pred. No. 6.3e-224;

Matches 510; Conservative 31; Mismatches 46; Indels 44; Gaps 4;

QY 1 MDFLLALVIVSSLYLQAAAEFGDGRWPRQIVSSIGLRYGGRIDCCGWARQSWGQCPFY 60  
 Db 1 MAVLLAAVLASSLYLQVAADFDGRWPRQIVSSIGLRYGGRIDCCGWARQSWGQCPFY 60

QY 61 VLKRIARIRCOLKAVCOIPKCKHGCIGPNKCKCHPGYACKTCIGVINEGGLKPRCKHR 105  
 Db 61 VLKRIARIRCOLKAVCOIPKCKHGCIGPNKCKCHPGYACKTCIGVINEGGLKPRCKHR 120

QY 106 -----VLNEGGLKPRCKHRCMNTYGSYKCYCLNGYMLMDGSCSSALTC 149  
 Db 121 PGPDPHQAATNVPDSHINCCGLKHPKCKHRCMNTYGSYKCYCLNGYMLMDGSCSSALTC 180

QY 150 CSMANCOYGVDDVWKGIRCCQSPSPGLQAPDGRVIFRQPSNDLFEIFEIE 209  
 Db 181 CSMANCOYGVDDVWKGIRCCQSPSPGLQAPDGRVIFRQPSNDLFEIFEIE 240

QY 210 YICKCHGKFDLMYIGGYQCHIDECSLGQYVTPKVMIPESGPVHVPKGNCTILKGDTON 269  
 Db 241 YICKCHGKFDLMYIGGYQCHIDECSLGQYVTPKVMIPESGPVHVPKGNCTILKGDTON 300

QY 270 QVYIPKVMIPESGPVHVPKGNCTILKGDTONNNWIPDVGSTWMPKPYIPPIITNRPST 329  
 Db 301 QVYIPKVMIPESGPVHVPKGNCTILKGDTONNNWIPDVGSTWMPKPYIPPIITNRPST 360

QY 330 KPTTPTTGLTTIAPAATSPGGITVDRVQTDQKPRGDFVIFRQPSNDLFEIFEIE 388  
 Db 361 KPTTPTTGLTTIAPAATSPGGITVDRVQTDQKPRGDFVIFRQPSNDLFEIFEIE 419

QY 389 NRVTDPQKPRGDFVIFRQPSNDLFEIFEIRGVSADEAADDPCVIVHSCNFDHGLCGW 448  
 Db 420 NRVTDPQKPRGDFVIFRQPSNDLFEIFEIRGVSADEAADDPCVIVHSCNFDHGLCGW 479

QY 449 IREKNDLHWEPTRDPAGGYLTVSAAKAPGKAAALWGRNGHGWQ 508  
 Db 480 IREKNDLHWEPTRDPAGGYLTVSAAKAPGKAAALWGRNGHGWQ 539

QY 509 LHSCTLOVFKKHGAIGAAALWGRNGHGWQ 568  
 Db 540 LHSCTLOVFKKHGAIGAAALWGRNGHGWQ 599

QY 569 DQVSLKRGRC 578  
 Db 600 DQVSLKRGRC 609

## RESULT 3

Q91V88 PRELIMINARY; PRT; 561 AA.

AC Q91V88;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DE FORM (NPHRONECTIN short isoform).  
 GN NPNT OR FORM OR NEPH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBL\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6;  
 RX MEDLINE=21551216; PubMed=11546798;  
 RA Morimura N., Tezuka Y., Watanabe N., Yasuda M., Miyatani S.,  
 RA Hozumi N., Tezuka K.;  
 RT "Molecular cloning of POEM, A novel adhesion molecule that interacts  
 with alpha8beta1 integrin.";  
 RL J. Biol. Chem. 276:42172-42181(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NIH SWISS; TISSUE=KIDNEY;  
 RX MEDLINE=21363579; PubMed=11470831;  
 RA Brandenberger R., Schmidt A., Linton J., Wang D., Backus C., Denda S.,  
 RA Muller U., Reichardt L.F.;  
 RT "Identification and characterization of a novel extracellular matrix  
 protein nephronectin that is associated with integrin alpha8beta1 in  
 the embryonic kidney.";  
 RL J. Cell Biol. 154:447-458(2001).  
 CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 DR EMBL: BAB59656; BAB9692.1; -;  
 DR EMBL: AY035898; AAK96010.1; -;  
 DR MGD: MGI:2148811; Npnt.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR000998; MAM\_domain.  
 DR Pfam: PF00008; EGF; 4.  
 DR Pfam: PF00629; MAM; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_3.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_3.  
 DR PROSITE: PS01187; EGF\_CA; UNKNOWN\_3.  
 DR PROSITE: PS50060; MAM\_2; 1.  
 KW Glycoprotein.  
 SQ SEQUENCE 561 AA; 61490 MW; 697ACAA0EE3F506 CRC64;  
  
 Query Match 85.4%; Score 2809.5; DB 11; Length 561;  
 Best local Similarity 85.3%; Pred. No. 7.4e-218;  
 Matches 494; Conservative 30; Mismatches 36; Indels 19; Gaps 3;  
  
 QY 1 MDFLLALVLSYLOAAAEFDGRWPQIVSSIGLCRYGGRIDCCWGWARQSGOCQPFY 60  
 DB 1 MAVLLAAVLAASSYLQVAAFDGRWPQIVSSIGLCRYGGRIDCCWGWARQSGOCQPFY 58  
  
 QY 61 VLQRQIARICQLKAVCPQCKHCEGICPNKCKCHPCYAGKTCIQVINECCGLKPRCKHR 120  
 DB 59 -----VCQPCQKHGECVGNPKCKCHPCFAGKTCNQDLNKGGLKPRCKHR 103  
  
 QY 121 CMNITYGKYCYCLNGYMLPDGSCSSALTCSMANCQYGCQVWKGQIRCOCPSPGLQIAPD 180  
 DB 104 CMNIFGSKYCYCLNGYMLPDGSCSSALSCSMANCQYGCQVWKGQVRCQCPSPGLQIAPD 163  
  
 QY 181 GRTCDVDGCATGRASCPFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGOY 240  
 DB 164 GRTCDVDGCATGRASCPFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGOY 223  
  
 QY 241 QCSSFARCYNVRGSKCKREGYOGDGLTCVYIPKVMIEPSGPIHVPKNGTILKGDGNG 300  
 DB 224 QCSSYARCYNIHSGYKCKQCRDGYEGDGLNCVYIPKVMIEPSGPIHVPKNGTILKGDGNG 283  
  
 QY 301 NNWIPDNGSTWPKTYIPIITNRPYTSKPTTRTPKPTPIPIPPPPPLPTELRL TPL 359  
 DB 284 ANRIPDAGSTRWPLKTYIPIITNRPYTSKPTTRTPKPTPIPIPPPPPLPTELRL TPL 343  
  
 QY 360 PPTTPERTGLTTIAPAASTPPGGITVDNRVQTDPOKPRGDVFIPIPPSNDLFEIFEIE 419  
 DB 344 PP-TPERPSTRPTIATSTTTTWTVDNRVQTDPOKPRGDVFIPIPPSNDLFEIFEIE 402  
  
 QY 420 RGVSAADFAKDDPGLVHSCNFDHGLCGWIREKNDLHWEP[IRDPAGQVLTLSAAKAPG 479  
 DB 403 RGVSADEEVKDDPGLLIHSCNFDHGLCGWIREKNDLHWETARDPAGQVLTLSAAKAPG 462

QY 480 GKAAARLVLPGLRLMSSDLCLSFRHKVTLGHSGLTGVFVRKKGACUAAALWGRNGHGWRQ 539  
 DB 463 GKAAARLVLRGLHLMHSHSDLCLSFRHKVTLGHSGLTGVFVRKKGACUAAALWGRNGHGWRQ 522  
  
 QY 540 TQITLRGADIKSVFVKGEKRRGHTGTEIGLDDVSLKKGHC 578  
 DB 523 TQITLRGADVKSVPFKGEKRRGHTGTEIGLDDVSLKKGRC 561  
  
 RESULT 4  
 Q91XL5 PRELIMINARY; PRT: 592 AA.  
 AC Q91XL5;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Nephronectin.  
 GN NPNT OR NEPH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBL\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6; TISSUE=KIDNEY;  
 RA Brandenberger R., Schmidt A., Linton J., Backus C., Wang D., Denda S.,  
 RA Muller U., Reichardt L.F.;  
 RT "Identification and Characterization of Nephronectin, a Novel ECM  
 Protein that is Associated with Integrin alpha8beta1 in the Embryonic  
 Kidney.";  
 RL J. Cell Biol. 0-0-0(2001).  
 CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 DR EMBL: AF397007; AAK84391.1; -;  
 DR MGD: MGI:2148811; Npnt.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR000998; MAM\_domain.  
 DR Pfam: PF00008; EGF; 4.  
 DR Pfam: PF00629; MAM; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_3.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_3.  
 DR PROSITE: PS01187; EGF\_CA; UNKNOWN\_3.  
 DR PROSITE: PS50060; MAM\_2; 1.  
 KW Glycoprotein.  
 SQ SEQUENCE 592 AA; 64944 MW; 21943B9077691396 CRC64;  
  
 Query Match 84.6%; Score 2784; DB 11; Length 592;  
 Best local Similarity 81.0%; Pred. No. 8.9e-216;  
 Matches 494; Conservative 30; Mismatches 36; Indels 50; Gaps 4;  
  
 QY 1 MDFLLALVLSYLOAAAEFDGRWPQIVSSIGLCRYGGRIDCCWGWARQSGOCQPFY 60  
 DB 1 MAVLLAAVLAASSYLQVAAFDGRWPQIVSSIGLCRYGGRIDCCWGWARQSGOCQPFY 58  
  
 QY 61 VLQRQIARICQLKAVCPQCKHCEGICPNKCKCHPCYAGKTCIQVINECCGLKPRCKHR 105  
 DB 59 -----VCQPCQKHGECVGNPKCKCHPCFAGKTCNQDLNKGGLKPRCKHR 103  
  
 QY 106 -----VLNECCGLKPRCKHRMNTVGSYKCYCLNGYMLPDGSCSSALT 149  
 DB 104 PLFQPPDHOATNVPISDLNECCGLKPRCKHRMNTVGSYKCYCLNGYMLPDGSCSSALT 163  
  
 QY 150 CSMANCQYGCQVWKGQIRCOCPSPGLQIAPDGRGTCDVDIDECATGRASCPFRQCVNTFGS 209  
 DB 164 CSMANCQYGCQVWKGQIRCOCPSPGLQIAPDGRGTCDVDIDECATGRASCPFRQCVNTFGS 223  
  
 QY 210 YICKCHKGFDLMYIGGKYQCHDIDECSLGOYCSSFARCYNVRGSKCKCKEYQGDGLT 269  
 DB 224 YICKCHKGFDLMYIGGKYQCHDIDECSLGOYCSSYARCYNIHSGYKCKQCRDGYEGDGLN 283

270 CYYIHKVMIPIISGPIHVIVKNGCTILKGDTGGNNMIPWGVSGTWPPKKPVPYIPPIITNRPRTS 329  
 IIIIIIIIIIIEEERNTTISKKDGQGHANRTPDAGSTRWPLKTPIPIPVITNRPTS 343  
 284 CVYIEKVMIIESGPILHMEERNGTISKDGQGHANRTPDAGSTRWPLKTPIPIPVITNRPTS 343  
 330 KPTTRPTPKPTPIPTPPPPPPLPTELR-TLPPTPTPERPTGLTTIAFAASTPPCGITVD 388  
 IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII 402  
 344 KYTTRTPINPTQQPTPPPPPPPPPPPTPEPRPTTLDP-TPERPSTRPTFLAVALSTTVIKVITVD 402  
 389 NRVTDPDKPRGDVFIPROPSNDLFEIFEIERGSVADEAKDDPGVLVIHSNFDHGICGW 448  
 IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII 462  
 403 NRIQTDPDKPRGDVFIPROPTNDLFEIFEIERGSVADEEVKDDPGILIHSCNFDHGICGW 462  
 449 IREKINDLHWKPIRIDPAGCOYLITVSAAKAPCKGAARLVLPICRLMHSGDMCLSPRHVKVG 508  
 IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII 522  
 463 IREKSDLIHWETARDPAAGCYLITVSNAKAPGKAARLVLRGLHLMESGDGLCSFRHKVKG 522  
 509 LHSGLQLQVVRKHGAIAGAALWGKRGNSHHWRQTQITLRGADTKSVYFKGEKERGHITGICL 568  
 IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIIIII 582  
 523 LHSGLQLQVVRKHGTHGAAIWGRNGCHGMHQQTQITLRGADVKSVIPFGKKRKRGHTGICL 582  
 QY 569 DQVSLKKKGHC 578 IIIIIIIIII I  
 DB 583 DQVSLKRGRC 592 IIIIIIIIII I

---

```

RESULT 5
ID QNY67
QNY67 PRELIMINARY: PRT; 554 AA.
AC QNY67;
DC 01-OCT-2000 (TREMBLrel, 15, Created)
DE 01-OCT-2000 (TREMBLrel, 15, Last sequence update)
DI 01-JUN-2002 (TREMBLrel, 21, Last annotation update)
DR Hypothetical 61.4 kDa protein.
DS Homo sapiens (human).
GN HMO.
OS Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=TERATOCARCINOMA/NEURON;
CA Franco H.;
RN Submitted (AUG-1999) to the EMBL/Genbank/DDAJ databases.
RI 121
RR SEQUENCE FROM N.A.
RC TISSUE=TERATOCARCINOMA/NEURON;
RX MEDLINE-20241927; PubMed-10777661;
XX Buchner G., Orfanelli U., Quaderi N., Bassi M.T., Andolfi G.;
RT Identification of a new EGF-repeat-containing gene from human Xp22:
XT A candidate for developmental disorders.*;
RL Genomics 65:16-23(2000)
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
EMBL; AJ245671; CAB92132.1; -.
DR HSSP; P00736; IAPC.
DR InterPro; IPR000152; Asx_hydroxyl,
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR000998; MAM_domain.
DR Pfam; PF00008; EGF_4.
DR Pfam; PF00629; MAM; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_like; 2.
DR SMART; SM00137; MAM; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS50060; MAM_2; 1.
KW calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
SQ Hypothetical protein; Repeat.
SK HYPOTHEICAL 554 AA; 61388 MW; D51923BF2A604101 CRC64;
SEQUENCE 554 AA; 61388 MW; 13.3%; Score 1096; DB 4; Length 554;

```

DR Pfam: PF00629; MAM\_1.  
DR SMART: SM00179; EGF\_CA; 3.  
DR SMART: SM00001; EGF\_like; 2.  
DR SMART: SM00137; MAM; 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01187; EGF\_CA; 3.  
DR PROSITE: PS00060; MAM\_2; 1.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 553 AA: 61314 MW: 2F55F167857DE50 CRC64;  
  
Query Match 33.2%; Score 1092.5; DB 4; Length 553;  
Best local Similarity 37.7%; Pred. No. 1.1e-79;  
Matches 215; Conservative 96; Mismatches 173; Indels 87; Gaps 14;  
  
QY 34 GLCRYGGRIDCCMGARQSWGQCQPFYVLRQRIARICQLKAVCPCKKHGECIGPNKCK 93  
Db 39 GVCHYGTKLACCYGWRNSKVC-----ATCEPCKEGECVGNPKCR 81  
  
QY 94 CHPGYAGTKCIQVINEGCLKPRCKHRCMNTYGSYKCYCLNGYMLMPDSCSALTCSMA 153  
Db 82 CFPGYTKCTQDVNECGMPPRCQHRVNTGSKYKFCLSGHMLLPDATCVNSRTCAI 141  
  
QY 154 NCQYGCVDVWVGQIRCOCPSPGLQAPDGRVCVDVDECATGRASCPRCQVNTFGSYCK 213  
Db 142 NCQYSCEDTEGPOCLCPSSGLRLAPNGRCLDIDECASGVICPYNRRCVNTFGSYCK 201  
  
QY 214 CHGFDLMYIGGYOCHDIDCSLGQYQCSFARCVNVRGYSYKCKEGYOGDGLTCVYI 273  
Db 202 CHIGFELQYISCRYDCIDINECTWDSHTCSHANCFTQSFCKCKQYKGNLRCSAI 261  
  
QY 274 PKVMEPSGPIHVPKGNGLTKGDTGNNNIPDVGSTWPPKTPYIPPIITNRPSTKPTT 333  
Db 262 PENSVK-----EVLRAFGTI-----KDRIKKLLAHRNKKKKA 294  
  
QY 334 R---PTPKPTPIPTPPPPPLPTLPTPTPTPTPTPTTGLTTIAPASIPPG----- 383  
Db 295 KIKNVTPEPTPTPT-----KYNLQPFNVE-----IVSRGNSHGSKKGNDEK 338  
  
QY 384 ---GITVDNRVQ---TDPQKRGDVFIPRQPSNDLFEIFEIERGVSADDEAKDDPGV 434  
Db 339 MKEGLEDKREKALKNDXERSLRGDVFPKVNAGEFGLLVORALKATSKLEHKDLNI 398  
  
QY 435 LVHSCNFDHGLCGWIREKNDLHWEPI RDPAAGQYLTVSAAPGKAARLVPLGLRLM 493  
Db 399 SV-DCSFNHGICDWKQREDQFQWNPADRONAIGFYMAVPALAGHKDKDIGRLKLLPDLQ 457  
  
QY 494 HSGDILCLSFRIKVTCLHSGTILQVFRKHGAHGAALWGRNGCHG--WRQTQITL-RGAD-I 549  
Db 458 POSNFCLLFDYRLAGDKVGRVFEVK--NSNNALANEKTTSEDEKWKTKIQLYOGTDTAT 515  
  
QY 550 KSVVFKGKRRGHTGEIGLDDVSLKKGHCSE 580  
Db 516 KSLIFAEKRGKCTGELAVDGVLLVSGLCPD 546  
  
RESULT 7  
Q9JJZ5 PRELIMINARY; PRT; 550 AA.  
AC Q9JJZ5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 61.5 kDa protein.  
GN EGF\_L6 OR W80.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Franco H.;

Submitted (AUG-1999) to the EMBL/GenBank/UDRJ databases.  
[2]  
SEQUENCE FROM N.A.  
RA MEDLINE=2041927; PubMed=10777661;  
RX Buchner G., Orfanelli U., Quaderi N., Bassi M.T., Andolfi G.;  
RT Identification of a new EGF-repeat-containing gene from human Xp22;  
RT Candidates for developmental disorders.;  
RI Genomics 65:16-23(2000).  
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
DR EMBL: AJ245672; CAB92138.1; --  
DR BSSP: P35555; 1EMN  
DR MGD; MG1:185859; Egf16.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001881; EGF\_CA.  
DR InterPro: IPR000998; MAM\_domain.  
DR Pfam: PF00008; EGF\_4.  
DR Pfam: PF00629; MAM\_1.  
DR SMART: SM00179; EGF\_CA; 3.  
DR SMART: SM00001; EGF\_like; 2.  
DR SMART: SM00137; MAM; 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01187; EGF\_CA; 3.  
DR PROSITE: PS00060; MAM\_2; 1.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;  
KW Hypothetical protein; Repeat.  
SQ SEQUENCE 550 AA: 61520 MW: DEF936325C9F31B3 CRC64;  
  
Query Match 33.0%; Score 1085; DB 11; Length 550;  
Best local Similarity 38.3%; Pred. No. 4.3e-79;  
Matches 217; Conservative 93; Mismatches 180; Indels 76; Gaps 13;  
  
QY 34 GLCRYGGRIDCCMGARQSWGQCQPFYVLRQRIARICQLKAVCPCKKHGECIGPNKCK 93  
Db 37 GVCOYGTMACCYGWRNKNAGVCE-----AMCEPCKEGECVGNPKCR 79  
  
QY 94 CHPGYAGTKCIQVINEGCLKPRCKHRCMNTYGSYKCYCLNGYMLMPDSCSALTCSMA 153  
Db 80 CFPGYTKCTQDVNECGVPRPCQHRVNTGSKYKFCLSGHMLLPDATCVNSRTCARL 139  
  
QY 154 NCQYGCVDVWVGQIRCOCPSPGLQAPDGRVCVDVDECATGRASCPRCQVNTFGSYCK 213  
Db 140 NCQYGCEDTEGPRCVCPSSGLRLAPNGRYCLDIDECASSKAVCPNRRCVNTFGSYCK 199  
  
QY 214 CHGFDLMYIGGYOCHDIDCSLGQYQCSFARCVNVRGYSYKCKEGYOGDGLTCVYI 273  
Db 200 CHIGFELKYIGRYDCVDINECALNTHPCSPHANCNLTGRSPKCKQYKGNGLQCSVI 259  
  
QY 274 PKVMEPSGPIHVPKGNGLTKGDTGNNNIPDVGSTWPPKTPYIPPIITNRPSTKPTT 333  
Db 260 PENSVK-----EVLTAFGTI-----KDRIKKLLAHRNKKKV 292  
  
QY 334 R---PTPKPTPIPTPPPPPLPTLPTPTPTPTPTTGLTTIAPASIPPG----- 380  
Db 293 KLMVTPRPASTRVP-----KYNLPSYSEEGYSGRNYDGEQKKEGKERLEE 342  
  
QY 381 PPGGIVDNRVQDTPQKRGDVFIPRQPSNDLFEIFEIERGVSADDEAKDDPGVLIHSON 440  
Db 343 EKGGKTLNRNVEQF-RTLRGDVFSPKVNEDLDLVYVORKEINSLKHKKDNLISV-DCS 400  
  
QY 441 FDHGLCGWIREKNDLHWEPI RDPAAGQYLTVSAAPGKAARLVPLGLRLMHISGDLG 499  
Db 401 FDLGVCMDKQDREDDFDWHPADRDNDVGYMAVPALAGHKKNIGRLKLLPLNLTQPSNFC 460  
  
QY 500 LSPRHKYTGILHSGTILQVFRKHGAHGAALW--GRNGCHGWROQTITL-RGAD-IKSVFVK 555  
Db 461 LLEFYRLAGDKVGRVFEVK--NSNNALANEETKNEDGRWRTKIQLYOGIDTKSVIFE 518  
  
QY 556 GEXRGTGHTGEIGLDDVSLKKGHCSEE 581  
Db 519 AERCKGKTGEIADVGVLLVSGLCPPD 544

```

RESULT 8
Q90FK6 PRELIMINARY: PRT: 558 AA.
AC Q90FK6:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2002 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 61.8 kDa protein (Eragrostis).
GN DKE2P564P2063.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Dusterthoef A., Lauber J., Mowes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL: AL117610; CAB56014.1; -.
DR BSSP: P00736; IAPQ.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR0000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR000998; MAM_domain.
DR Pfam: PF00008; EGF_4.
DR Pfam: PF00629; MAM_1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_like; 2.
DR SMART: SM00137; MAM; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 2.
DR PROSITE: PS00060; MAM_2; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Hypothetical protein; Repeat.
FT NON_TER
SQ SEQUENCE 558 AA; 61828 MW; AA38D7DCE402BFA3 CRC64;

Query Match 32.4%; Score 1066.5; DB 4; Length 558;
Best local Similarity 37.5%; Pred. No. 1.4e-77;
Matches 214; Conservative 95; Mismatches 175; Indels 87; Gaps 15;

QY 34 GLCRYGGRIICGWCWARSWQCCPFYVLRQRIARICQLAAVCPCKHCEGCPNCK 93
DB 44 GVCHGTKLACCYGWRNSKVC-----ATCKGCKFGCGVCPNCKR 86

QY 94 CHPYAGKTCIGVLENEGLKPRCKIRCMNTYGSYCYCLNGYMLPDPGSSSALTCMA 153
DB 87 CLPGYTGKTCSDVNEGCKPRPCQIRCVNTHGSYKCFCLSGHMLPDPATCVSRTCA 146

QY 154 NCQYCCDVVKCOIRCOCHSPGLQIAPGRTCVVDPCATGRASCPKCVNTFGSYCK 213
DB 147 NCQYSCDETEGPQCLPSSGLRLAPNGRGLDIDECASGRVICPYNRKCVNTFGSYCK 206

QY 214 CHKGFDLAVIGKCYCHDIDECISLGOYCCSSFARCYNVRGSKCKEGYCGDGLTCVY 273
DB 207 CHIGFELQYISGRVYCIDINCTMISHSTSSHANCFTNGSKCKCKGKNGKNCAL 266

QY 274 PKVMIEPSGPIHVPGKNGTILKGTGNNNWIPIVGVSTWWPKTPYIPPIINRPTSKPT 333
DB 267 PENSVK-----EVLRAPTI-----KRIKLLAIKNSMKKKA 299

QY 334 R---PTPKPTPIPTPPPPPLPTELRPLDPTTPPTTGLTITAPAASDPG----- 383
DB 300 KIKNVIPEPTPTPT-----KVNLPQFNVE---GT---VSRGNSHGKKGNEEK 343

QY 384 ---GTVDNRVG-----TDCKPKRGDVFIPRQPSNDLFEIFEIERGVSADDAKDPGV 434
DB 344 MKEGLEDKRRKPKALKNIDIERNSKGVDFPKVNFVAFAGEGILIVORKALTSKLEHKDLNI 403

```

```

QY 435 LVHSONFDHGLCGWIRKDNIDLHWRPILRIDACQYLTVSAAKAKGCKAARLVLMGRKM 494
DB 404 SV-DCSENHGICDKWQDREDFDNPALRONAIGFYMAVPALAGIKKIKLKLPLQLQ 462

QY 494 HSGDILCLSFRRHKVTGLHSGTLOVFKRHGAHGAALMGKNGCHG--WRGTQITLKGAD 1549
DB 463 POSNFCSLFDYRIAGDKVKIKRVFVK--NSNNALAEKTTSEDFKWKIKGKILQIYQNTLAI 520

QY 550 KSVVPKREKRGITGEIGLDVSLKKKGICSE 580
DB 521 KSTIPFAHKGKGTGKTGTAVDGCVLLVSGIYD 551

RESULT 9
Q8WYG3 PRELIMINARY: PRT: 474 AA.
AC Q8WYG3:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 53.1 kDa protein.
GN P6648.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.X., Wan D.F., Zhao X.F., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li B.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF193055; AAG22483.1; -.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR0000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR000998; MAM_domain.
DR Pfam: PF00008; EGF_4.
DR Pfam: PF00629; MAM_1.
DR SMART: SM00181; EGF_4.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00137; MAM; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_3.
DR PROSITE: PS00060; MAM_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 53144 MW; 5F116F0DF91AC8BD CRC64;

Query Match 27.2%; Score 895; DB 4; Length 474;
Best local Similarity 46.5%; Pred. No. 7e-64;
Matches 184; Conservative 89; Mismatches 161; Indels 70; Gaps 13;

QY 101 KTCIQVLENEGLKPRCKIRCMNTYGSYCYCLNGYMLPDPGSSSALTCMAWQVGT 160
DB 10 ETCSDVNEGCKPRPCQIRCVNTHGSYKCFCLSGHMLPDPATCVSRTCAVINQTSCE 69

QY 161 VVKQIRCOCHSPGLQIAPGRTCVVDPCATGRASCPKCVNTFGSYCKCKHGFGL 220
DB 70 DTEGPQCLPSSGLRLAPNGRGLDIDECASGRVICPYNRKCVNTFGSYCKKHGFEL 129

QY 221 MYIGKCYCHDIDECISLGOYCCSSFARCYNVRGSKCKEGYCGDGLTCVYIPKVMIEP 280
DB 130 QYISGRVYCIDINCTMISHSTSSHANCFTNGSKCKCKGKNGKNCALPENSVK 188

QY 281 SGPIHVPGKNGTILKGTGNNNWIPIVGVSTWWPKTPYIPPIINRPTSKPTTKE 437
DB 189 ----EVLRAPTI-----KRIKLLAIKNSMKKKA 222

QY 338 KPTPIPTPPPPPLPTELRPLDPTTPPTTGLTITAPAASDPG 487

```

```
Db 223 EPTPTPT-----KVNQLQFNVE-----IVSRGNSHGKKGNEKKMKGLD 266
Qy 388 DNRVQ-----TDPKPRGVDIFIPQPSNDLFEIFEIRGYSADDEAKDDPGVLVHSCNF 441
Db 267 EKREFKALKNDIEKSLRGDWFPPKVNACPGFGLTQVKALTSKLEHKDLNISV-DCSF 325
Qy 442 DHGLCCWIREKNDLWHEPI-RDAGGQYLTVSAAKAPGCKAARLVLPICRIHMSGDLC 500
Db 326 NHGCDMKQDREDDFDMNPADRDNAIGFYMAVPALAGHKDDIGELKLLDPLQPSNFC 385
Qy 501 SFRHKVTGLHSGTQLQVFKKCHGAALWGRNGCHG--WROTQTLT-RGAD-IKSVVFKG 556
Db 386 LPDYRLAGDKVKLRVPVK--NSNLAWEKTTSEDEKWKTKTLQIQGTDTAKSIIFEA 443
Qy 557 EKRRGHTGEIGLDDVSLKKHCSE 580
Db 444 ERGKGTGEIAVDGVLIVSGLCPT 467

RESULT 10
Q96JP8 PRELIMINARY; PRT; 2809 AA.
ID AC Q96JP8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibrillin3.
GN KIAA1776.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21245130; PubMed=11347906;
RA Nadase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
DR EMBL; AB053450; BAB47408.1; -
DR InterPro; IPR000152; ASX_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 45.
DR Pfam; PF00683; TB; 9.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_41.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_36.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_40.
SQ SEQUENCE 2809 AA; 300323 MW; 20C04CC006C0161F CRC64;

Query Match 12.3%; Score 405.5; DR 4; Length 2809;
Best Local Similarity 36.6%; Pred. No. 1.5e-23;
Matches 94; Conservative 26; Mismatches 98; Indels 39; Gaps 13;

Qy 35 LCRVG-----GRIDCCGWARQSGWQCQPFYVLKRIARI-----RCQLKAVCQPRCKHG 84
Db 1081 LCRGTCTNTGSKYC-----QCPCHETAKTACEDIDECSLS-----DGLCPHG 1127
Qy 85 ECGT---IGNKCKCHPGYAG---RTCIQVLNCEGLKPRCKHRCMNTYSGKYCYCLANGYM 137
Db 1128 OCNVNIGAFQCSCHAGPOSTPDQGCVDI-NPCRVNGCCDVHCINTGSKYKSCGGQYS 1186
Qy 138 LMPDG-SCSSAITC---SMANQYQ-CDVVVKQI-RCQCPSPGLIQLAPDCHTCVDVDECATG 193
Db 1187 LMPDGRACADYDECEENPRVCDQGHCTNMPGGRHCLC-YDGFMATPDMDRTCDVDVDECDLN 1245
Qy 194 NASCPFRPQCNTFGSYLCKCHKFDLMYIGKYQCHDIDFCSLGOYQCSSFARCYNVRG 253
```

```
Db 1246 PHIC-LHGDCNTKGSFVCHCQLQY--MVRKGAIGSDVDCHVGGHNCUSHASCLNIPC 1302
Qy 254 SYKCKCKEGYGGDLTC 270
Db 1303 SFSCRLCPGWVGDFEC 1319

RESULT 11
Q99K58 PRELIMINARY; PRT; 1174 AA.
ID AC Q99K58;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DI 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to fibulin 2.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005443; AAH05443.1; -
DR HSP; P00736; IAPQ.
DR MGD; MGI:95488; Fbln2.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; ASX_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_CA.
DR Pfam; PF01821; ANATO; 2.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00179; EGF_CA; 9.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 9.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 1174 AA; 126460 MW; 8D628AC710FBA688 CRC64;

Query Match 12.2%; Score 402; DB 11; Length 1174;
Best Local Similarity 38.9%; Pred. No. 9.9e-24;
Matches 88; Conservative 27; Mismatches 86; Indels 25; Gaps 12;

Qy 80 RCKHCE-CL---GPNKCKCHPGYA----GKTCIQVLNFCGLKLP-RPCKHRCMNTYSGSKC 130
Db 858 RCCEGOLCYNLPGSYRCDCPKPGFQDFAFGRTCIDV-NECWVSPGLAQHITCENTPGSYRC 916
Qy 131 YCLNGYMLMPDGS-CSSALTCSMANQYQCDVVKQIRCCPSPGLAPDGRTCVDVDE 189
Db 917 SCAAGPILADGKHCEVDNCETRRCQECANIVGSCQVC-RCQYGLADGHTCTDIDE 975
Qy 190 CATGRASCPFRQCNTFGSYICK-KHGFDMYIGRKYQCHDIDECSLGQYQCSSFARC 248
Db 976 CAQAGILCTFR-CVNVPGSYQACPEQGYTMANG--RSSCKDLDECALGTHNCSAEATC 1032
Qy 249 YNVKSGSKC---CKCKEY-----QGDLTCVYIPKVMIEPSGPIH 285
Db 1033 HNIQGSFRLCFDCCPPNVYRVSETKCRITTCQDITECQTSPARITH 1078

RESULT 12
Q88840 PRELIMINARY; PRT; 3857 AA.
ID AC Q88840;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DI 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Mutant fibrillin-1.
```

GN FBNI.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID-10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-B10.D2;  
 RX MEDLINE-98069008; PubMed-9405944;  
 RA Rona C.A., Murai C., Casares S., Kasturi K., Nishimura H., Honjo T.,  
 RA Matsuda F.,  
 RT "Structure of the mutant fibrillin-1 gene in the tight skin (TSK)  
 mouse.";  
 RL DNA Res. 4: 267-271(1997).  
 DR FMBI: AF007248; AAC62317.1;  
 DR HSSP: P35555; IADJ.  
 DR MGI: MGI-95489; Fbni.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR002557; Chitin\_bind\_Pera.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR002212; Fibril-assoc.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR Pfam: PF00008; EGF; 64.  
 DR Pfam: PF00683; TB; 12.  
 DR SMART: SM00494; ChIBD2; 2.  
 DR SMART: SM00179; EGF\_CA; 60.  
 DR SMART: SM00001; EGF\_Like; 4.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 61.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_2.  
 DR PROSITE: PS01186; EGF\_2; 50.  
 DR PROSITE: PS01187; EGF\_CA; 61.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat..  
 SQ SEQUENCE 3857 AA; 418301 MW; 5BC0618BC527E04C CRC64;  
 Query Match 12.2%; Score 400.5; DB 11; Length 3857;  
 Best Local Similarity 31.7%; Pred. No. 5.4e-23;  
 Matches 104; Conservative 33; Mismatches 104; Indels 87; Gaps 15;  
 QY 35 LCRVY-----GRIDCCGWAHSGOCQPPYVIRIARI-----RCOLKAVCQPRCKHG 84  
 DB 1125 LCRGICHTGSEYRC-----ECPGHQLSPNISACIDINECELSA---NLCPHG 1171  
 QY 85 EC---IGPNKCKCHPGYAGK-----TCIOVLNPGCLKPRCKHRCMNTYGSYKCYCLNGYM 137  
 DB 1172 RCVNIIGKYQCACNIGYHPITDRIKPCVDI-DKCSIMNGCGCTFTNSDGSYKSCQPCFA 1230  
 QY 138 LMPD-GSCSSALTC--SMANQYQ-CDVVKGQIRCCQSPGLQAPDGRCTVDV----- 187  
 DB 1231 LMPDQSCDTIDPCRDNPNICDGGQCTNIPGHRGIC-YDGFMASEDMDKTCVDVNFCDLN 1289  
 QY 188 -----DKCATGRASCPFRFCVNTFGSYI 211  
 DB 1290 PNICLSGTGTCNTKGSFICIDMGYSKKKGTGCTDINECEICAINCGRIHVCNTNTAGSFK 1349  
 QY 212 CKCHRGFDIMYIGKYGQCHDIDECISIGQYQCSFARCYNVKSGYKCKEGYQCGILTCV 271  
 DB 1350 CSCSPG---WIGDCIKCTIDECNSGTHMCSQAHKCNKNGSYKICLCKDYGITGIGFTCT 1405  
 QY 272 YPKVMIEPSGPIIYKNGITLKGDGTG 299  
 DB 1406 DLD-----ECSENIINI-CGNGCINAPG 1428  
 RESULT 13  
 Q9WUH9 PRELIMINARY; PRT: 2906 AA.  
 ID Q9WUH9  
 AC Q9WUH9  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Fibrillin-2.

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID-10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-99350241; PubMed-10419698;  
 RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashihara N., Wallner K.,  
 RA Kanwar Y.S.,  
 RT "Cloning of rat fibrillin-2 cDNA and its role in branching  
 morphogenesis of embryonic lung.";  
 RL Dev. Biol. 212:229-242(1999).  
 DR EMBL: AF150600; AAD34439.1;  
 DR HSSP: P35555; IEMN.  
 DR InterPro: IPR002086; Aldehyde\_dehydr.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000561; EGF-Like.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR002212; Fibril-assoc.  
 DR Pfam: PF00008; EGF; 46.  
 DR Pfam: PF00683; TB; 9.  
 DR PRINTS: PR03010; RGFHLOOD.  
 DR SMART: SM00179; EGF\_CA; 42.  
 DR SMART: SM00001; EGF\_Like; 4.  
 DR PROSITE: PS00070; ALDEHYDE\_DEHYDR\_CYS; UNKNOWN\_1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 43.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_2.  
 DR PROSITE: PS01186; EGF\_2; 36.  
 DR PROSITE: PS01187; EGF\_CA; 43.  
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat..  
 SQ SEQUENCE 2906 AA; 313371 MW; 9EF64E727044EF58 CRC64;  
 Query Match 12.0%; Score 396; DB 11; Length 2906;  
 Best Local Similarity 39.1%; Pred. No. 8.9e-23;  
 Matches 79; Conservative 33; Mismatches 74; Indels 16; Gaps 9;  
 QY 81 CKHGEK---IGPNKCKCHPGYAGK-----KTCTQVLNPGCLKPRCKHRCMNTYGSYKCYCL 134  
 DB 1203 CRNCKCVNMICTGYCCSCNPGYQATPDHGGSDI-DECMINNGGFTDCTNSRCSYKSCS 1261  
 QY 134 NGYMLMPDG-SCSSALTC--SMANQYQ-CDVVKGQIRCCQSPGLQAPDGRCTVDVDE 189  
 DB 1262 EGYALIMPGRRSCADIDECENNPDCDGGQCTNIPGHRGIC-YDGFMASEDMDKTCVDVNE 1420  
 QY 190 CATGRASCPFRFCVNTFGSYICKHKGFDIMYIGKYGQCHDIDECISIGQYQCSFARCY 249  
 DB 1321 CDLPNITC-MFEGECNTKGSFICHCQLQYSVK--KGATGCTDVIDEIEIGAINILMIASCL 1477  
 QY 250 NVKSGYKCKEGYQCGDLTCV 271  
 DB 1378 NVPGSFKSCHEGWGNGICKI 1399  
 RESULT 14  
 Q9WUH8 PRELIMINARY; PRT: 2872 AA.  
 ID Q9WUH8  
 AC Q9WUH8  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Fibrillin-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID-10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-99032689; PubMed-9815129;  
 RA Kanwar Y.S., Ota K., Yang Q., Kumar A., Wada J., Kashihara N.,  
 RA Peterson D.R.,  
 RT "Isolation of rat fibrillin-1 cDNA and its relevance in metaphoric  
 development.";



```

RL Am. J. Physiol. 275:F710-F723(1998).
RN [2]
RA SEQUENCE FROM N.A.
RP Kanwar Y.S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF135059; AAD34438.1;
DR FSSP: P35555; 1APJ.;
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; Fibril-assoc.
DR InterPro: IPR000822; ZnF_C2H2.
DR Pfam: PF00008; EGF; 45.
DR Pfam: PF00683; TB; 9.
DR SMART: SM00179; EGF_CA; 41.
DR SMART: SM00001; EGF_like; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; 42.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; 38.
DR PROSITE: PS01187; EGF_CA; 41.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW KQ
SQ SEQUENCE: 2872 AA; 312069 MW; 0C4F7F3B87A80280 CRC64;

Query Match 12.0%; Score 395.5; DB 11; Length 2872;
Best Local Similarity 31.4%; Pred. No. 9.6e-23;
Matches 103; Conservative 33; Mismatches 105; Indels 87; Gaps 15;

QY 35 LCRYG-----GRIDCWGHARQSWGQCQPPYVLRORIARI----RCQLKAVCQPRCKHG 84
DB 1124 LCRGGICHNTEGYSYR-----ECPSGHQSPNISACIDINEGELSA---NLCPHG 1170
QY 85 EC---ICPNKCKCHPGYAGK---TCIQVINEGCLKPRCKHRCMNITGYSKYCYCLNGYM 137
DB 1171 KCVNLGKYECACNPQYHPHTRDLFCVDI--DECSIMNGGCTCTNSDGYECSQCPQFA 1229
QY 138 LMPD-GSCSSALTC--SMANCQYG-CDVVKQIRCCQSPGLQLAPDRTCDVY-----187
DB 1230 LMPDORSCTDIDCEDNPNICDGGCQCTNIPGYRCLC-YDGFMASEDMKTCVDVNFCDLN 1288
QY 188 -----DECATGRASCPFRQCVNTFGSYI 211
DB 1289 PNCLSGTCENTKSGFTCHCDMGYSGKKGTGTDTINECEIAGNCRGHAVCTNTAGSPK 1348
QY 212 CKCHKGFDLWVIGKYOCYCHDIECSLQYQYCSSFARCVNVRGYSKCKCKEYQCGDGLTCV 271
DB 1349 CSCSPG---WICDGIKCTDLDECSNTHMCSHQADCKNTGMSYRCLCKDGYTGSGFTCT 1404
QY 272 YIPKVMTEPSGPIHVPKGNITLKGDTG 299
DB 1405 DLD----KCSNLNL-SGNGOCLNAPAG 1427

RESULT 15
Q9DBE2
ID Q9DBE2 PRELIMINARY; PRT; 608 AA.
AC
DT 01-JUN-2001 (TREMBLrel, 17, Created)
DT 01-JUN-2001 (TREMBLrel, 17, last sequence update)
DT 01-JUN-2002 (TREMBLrel, 21, last annotation update)
DE 1300015804Rik protein.
GN 1300015804Rik.
OS Mus musculus (Mouse).
OC Pukayota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RA SEQUENCE FROM N.A.
RP STRAIN-C57H1/6J; TISSUE=LIVER;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

```

Wed Apr 23 13:51:06 2003

us-09-635-949-34.rspt

Page 10

Search completed: April 23, 2003, 13:03:12  
Job time : 101 secs

---